



Excess primer degradation by Exo I improves the preparation of 3' cDNA ligation-based sequencing libraries

Enroth, Christel Hougård; Fehler, Annaleigh Ohrt; Poulsen, Line Dahl; Vinther, Jeppe

Published in:
BioTechniques

DOI:
[10.2144/btn-2018-0178](https://doi.org/10.2144/btn-2018-0178)

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Document license:
[CC BY-NC-ND](#)

Citation for published version (APA):
Enroth, C. H., Fehler, A. O., Poulsen, L. D., & Vinther, J. (2019). Excess primer degradation by Exo I improves the preparation of 3' cDNA ligation-based sequencing libraries. *BioTechniques*, 67(3), 110-116.
<https://doi.org/10.2144/btn-2018-0178>

Excess primer degradation by Exo I improves the preparation of 3' cDNA ligation-based sequencing libraries

Christel Hougård Enroth^{†,1}, Annaleigh Ohrt Fehler^{†,1}, Line Dahl Poulsen¹ & Jeppe Vinther^{*,1}

ABSTRACT

RNA sequencing library construction using single-stranded ligation of a DNA adapter to 3' ends of cDNAs often produces primer–adapter byproducts, which compete with cDNA–adapter ligation products during library amplification and, therefore, reduces the number of informative sequencing reads. We find that *Escherichia coli* Exo I digestion efficiently and selectively removes surplus reverse transcription primer and thereby reduces the primer–adapter product contamination in 3' cDNA ligation-based sequencing libraries, including small RNA libraries, which are typically similar in size to the primer–adapter products. We further demonstrate that Exo I treatment does not lead to trimming of the cDNA 3' end when duplexed with the RNA template. Exo I digestion is easy to perform and implement in other protocols and could facilitate a more widespread use of 3' cDNA ligation for sequencing-based applications.

METHOD SUMMARY

For 3' cDNA ligation-based RNA sequencing libraries, we demonstrate that Exo I treatment after reverse transcription reduces primer–adapter contamination without causing trimming of the cDNA 3' end. Our approach makes 3' cDNA adapter ligation more broadly applicable to library preparation, including small RNA-seq libraries, which typically are impossible to separate from adapter–primer contamination based on size selection.

KEYWORDS

cDNA ligation • Exonuclease I • RNA-seq • sequencing library preparation

¹Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark;

*Author for correspondence: jvinther@bio.ku.dk;

[†]These authors contributed equally to this work

BioTechniques 67: 110–116 (September 2019) 10.2144/btn-2018-0178

Massive parallel sequencing is a useful tool in transcriptomics and genomics and has been applied to the study of a multitude of different biological research questions over the last decade [1,2]. Sequencing of cDNA is now being used for gene expression analysis and for many other applications such as the study of small RNAs, protein–RNA interactions, RNA–RNA interactions, RNA structures and ribosome profiling.

Most sequencing methods, including the Illumina platform, require addition of adapter sequences to the ends of the DNA fragments [3] and many different strategies for adding these exist. For sequencing of double-stranded DNA, adapters can be added either by fragmentation of the DNA followed by ligation [4] or by transposon-based introduction of the adapters [5]. These strategies can be used to prepare libraries from cDNA-derived double-stranded DNA; however, for some applications this is not convenient or possible. Instead, adapter sequences can be ligated to the ends of RNA before reverse transcription, included in the reverse transcription primer or ligated to the cDNA 3' end. Alternatively, the terminal transferase activity of M-MLV reverse transcriptase can be used in combination with its ability to switch template to add a 3' cDNA adapter sequence during reverse transcription [6], or reverse transcription can be terminated stochastically with 3' azido labeled dideoxynucleotides followed by click chemistry-mediated adapter addition [7].

One of the most straightforward strategies for constructing cDNA sequencing libraries is to use a primer with an adapter overhang for reverse transcription and subsequently ligate a DNA adaptor to the 3' end of the cDNA

with either CircLigase or T4 DNA ligase. This 3' cDNA ligation strategy has successfully been used for RNA-Seq [8,9] and sequencing of transcriptome-wide RNA structure probing [10–13]. The 3' cDNA ligation reaction is known to be biased, but the use of T4 ligase significantly reduces the bias [13], meaning that 3' cDNA ligation is an attractive strategy for sequencing library preparation. However, widespread use has been limited by primer–adapter contamination arising from the adapter ligating to the 3' end of the reverse transcription primer. During PCR amplification of the sequencing library, the primer–adapter product is amplified more efficiently than the intended cDNA product because of its short length. If a large amount of RNA starting material is used, it is typically possible to produce cDNA sequencing libraries with low primer–adapter product contamination, while lower amounts of starting material produce libraries with considerable amounts of primer–adapter product, which will reduce the number of informative sequence reads obtained from the sequencing [3,13–15].

To avoid contaminating primer–adapter products in 3' cDNA ligation-based libraries, several different strategies can be applied. First, an increased amount of RNA starting material can in some cases minimize primer adapter contamination. Second, a reduction in the amount of primer used in the reverse transcription reaction reduces primer–adapter contamination, but the reduction also affects the efficiency of cDNA synthesis, meaning that this strategy requires optimization and may not work in all cases. Third, in many applications, the cDNAs of interest are longer than the primer–adapter, which makes ►

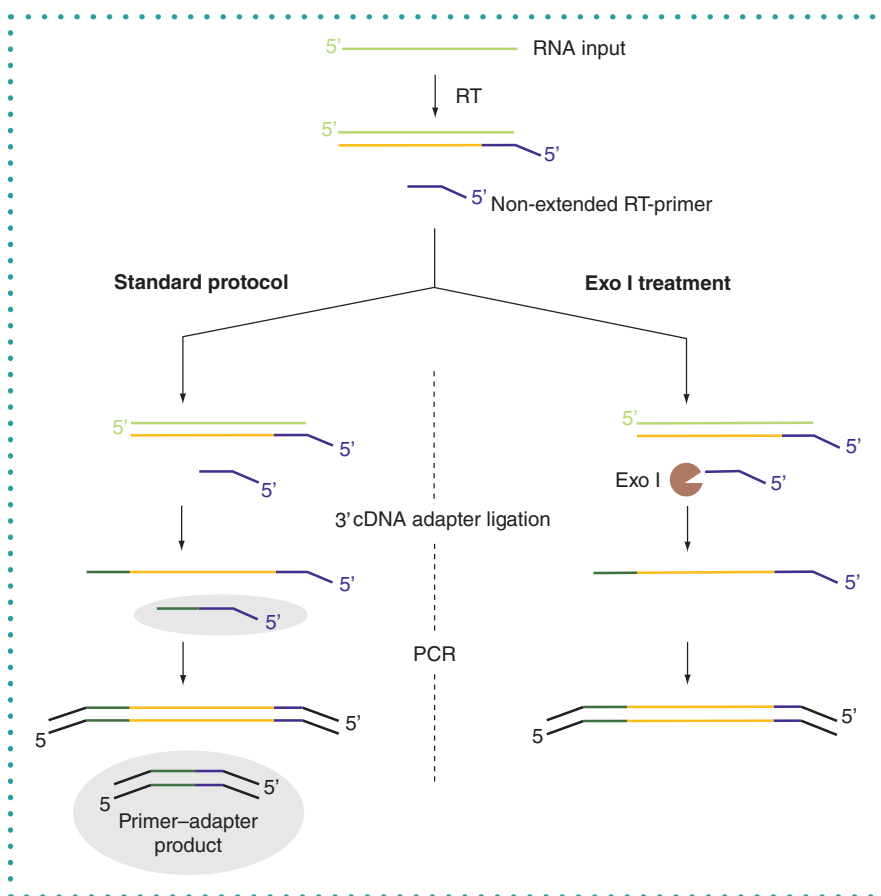


Figure 1. Schematic representation of the Exo I method. After reverse transcription of RNA to cDNA, surplus RT-primer will remain in the solution. In the standard protocol, the surplus RT-primer ligates to the adapter during 3' cDNA adapter ligation, thereby causing primer-adaptor contamination in the standard protocol (grey background), which is amplified in the PCR and ultimately reduces the number of informative sequenced reads. The Exo I treatment protocol uses Exo I to remove surplus RT-primer after RT, retaining only the cDNA/RNA duplex of interest at the 3' cDNA adapter ligation step to increase the proportion of informative sequenced reads in the sequencing library. RT: Reverse transcription.

optimization of the CLIP-Seq protocol, Exo I and alkaline phosphatase treatment were introduced to remove reverse transcription-primers before ligation of an adapter to the 3' end of cDNA with T4 RNA ligase [24]. Finally, in a recent RNA-seq protocol, Exo I was used together with selection on streptavidin beads to remove surplus reverse transcription primer [25]. None of these studies included data on the effect of Exo I treatment in library preparation or investigated if inclusion of Exo I in their protocol caused trimming of the 3' end of the cDNA.

Here, we demonstrate that Exo I treatment of reverse transcription reactions before cDNA ligation dramatically reduces the amount of primer-adaptor contamination in RNA-seq libraries, including small RNA sequencing libraries, and does not result in 3' end trimming of the cDNA.

MATERIALS & METHODS

RNA purification

E. coli growth conditions and RNA extraction are described in detail in the Supplementary methods. For small RNA library preparation, an equal mixture of five chemically synthesized 5' phosphorylated and HPLC-purified RNA oligonucleotides (Eurofins) with a mutated version of the let-7 miRNAs (mut-let-7a, b, c, e, f) (Supplementary Table 1) sequence were used.

Fragmentation

For fragmentation, 560 ng *E. coli* RNA was incubated in 50 mM Tris-HCl (pH 8) and 5 mM MgCl₂ at 95°C for 3 min and 20 s before placing on ice and adding EDTA to a final concentration of 10 mM in a final volume of 70 µl.

RNA & cDNA purifications between enzymatic reactions

For *E. coli* RNA-seq libraries, fragmented RNA was purified using Agencourt RNAClean XP beads. After subsequent steps (reverse transcription, Exo I digestion, adapter ligation and PCR), DNA was purified with Agencourt AMPure XP beads (Beckman Coulter, cat. no. A63987 and A63881) using 1 volume of the nucleic acid solution to 1.8 volume of beads according to manufacturer's protocol. The size-selected sample was subjected to an additional round of Ampure purification. For preparation of microRNA libraries, samples were purified after all steps (polyadenyl- ▶

▶ it possible to reduce contamination by size selection of the cDNA or the final sequencing library. However, the size selection step is time consuming and requires that sufficient cDNA extension products are present in the library. In addition, for libraries produced from small RNAs, such as miRNAs, it is difficult to size-separate the extended cDNA from the primer-adaptor product [16]. Fourth, inclusion of biotin-labeled nucleotides in the reverse transcription reaction allows streptavidin-based bead purification of cDNA products from the reverse transcription primer prior to ligation [13]. This strategy is, although useful for many applications, inefficient for shorter cDNAs and cannot be implemented for applications that require biotin labelling, such as RNA structure probing with N3-NAI [17], CAGE [18] or the SHAPES method [19].

For many years, the ability of *Escherichia coli* Exo I to catalyze degradation of single-stranded DNA without targeting double-stranded DNA [20] has been exploited for cleanup of PCR products. Excess primer can be degraded by Exo I and the resulting dNTPs dephosphorylated by alkaline phosphatase, thereby preparing PCR products for downstream applications [21]. While the lack of Exo I activity on DNA duplexes is well characterized [20], Exo I's activity on DNA in RNA-DNA duplexes has to our knowledge not been formally investigated and there are only a few examples of Exo I being used to remove reverse transcription primers after reverse transcription [22–25]. In two similar single-cell RNA-seq protocols, reverse-transcribed mRNA was treated with Exo I before poly-A tailing of the 3' cDNA end in the subsequent step [22,23]. Likewise, in a recent

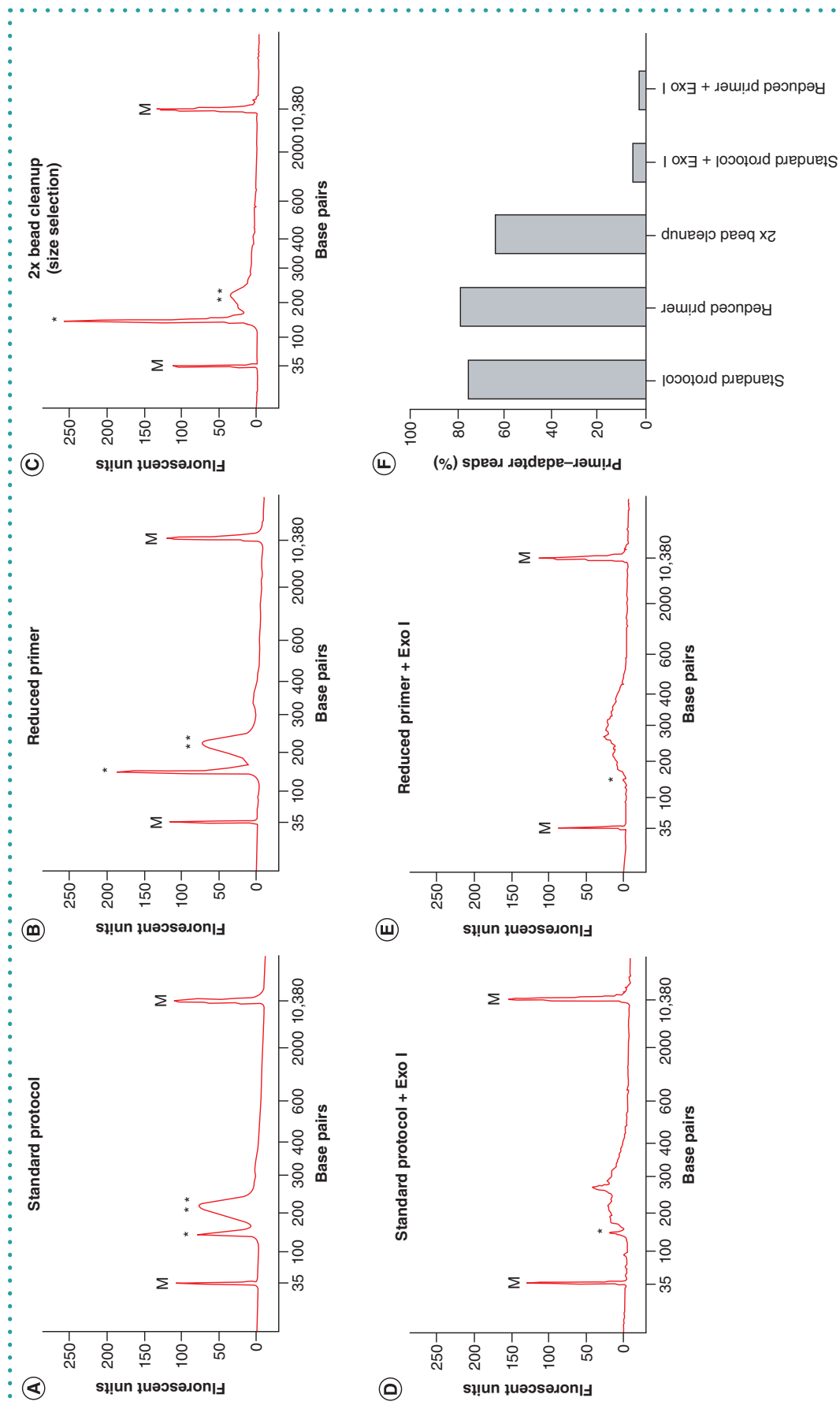


Figure 2. Bioanalyzer electropherogram traces of final *Escherichia coli* RNA sequencing libraries. (A) RNA sequencing library prepared with a standard protocol. (B) Standard protocol, but with a 5x reduction in amount of reverse transcription primer. (C) Standard protocol, but with size selection using additional Ampure XP bead cleanup following reverse transcription. (D) Standard protocol, but treated with Exo I following reverse transcription to remove leftover primer. (E) As (B), but treated with Exo I. (F) Percentage of primer-adaptor reads in the five libraries corresponding to percentage of read pairs that were 22 nt after trimming adapter sequences (see supplemental material and methods).

*Primer adapter peak (147 bp).

**Primer adapter concatemer.

M: Standard lower or upper size marker used for size estimation.

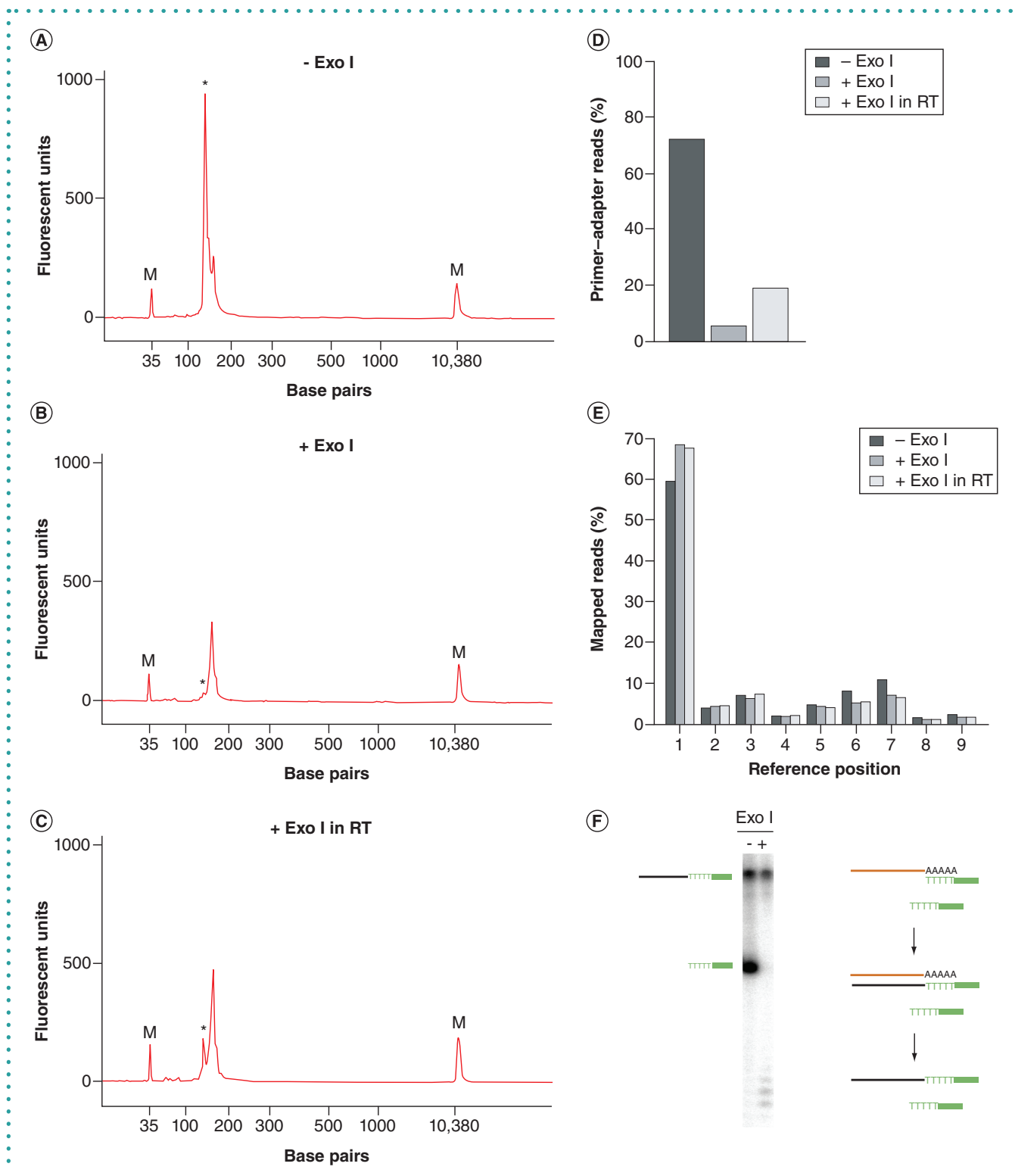


Figure 3. Bioanalyzer electropherogram traces of miRNA sequencing libraries. (A) MicroRNA sequencing library prepared without Exo I treatment. M: Standard lower or upper size marker used for size estimation. (B) MicroRNA sequencing library treated with Exo I in the Exo I buffer. (C) MicroRNA sequencing library treated with Exo I in the RT buffer. (D) Percentage of primer-adaptor reads in the three libraries (A–C) corresponding to processed reads that were too short to incorporate miRNA sequences (see Supplementary materials and methods). (E) Distribution of alignment start positions from in the reference miRNA sequences of the three libraries (A–C). (F) Denaturing acrylamide gel of 5' end ³²P-labeled poly-dT primer (Supplementary Table 1) in RT of polyadenylated miRNAs, treated with Exo I (+) and not treated with Exo I (–). The position of full-length cDNA (61 nt) and the non-elongated poly-dT primer (41 nt) in the gel is indicated.

*Primer-adaptor peak (147 bp).

M: Standard lower or upper size marker used for size estimation; RT: Reverse transcription.

► ation, reverse transcription, Exo I digestion, adapter ligation and PCR) by phenol-chloroform extraction and ethanol precipitation, with the exception of one sample, in which the reverse transcription reaction was used directly for Exo I treatment without prior purification.

Polyadenylation

Samples containing 510 ng miRNA oligonucleotide were polyadenylated with 5 U *E. coli* Poly(A) polymerase (New England Biolabs, cat. # M0276) in 1× *E. coli* Poly(A) polymerase buffer and 1 mM ATP in a total volume of 20 µl and incubated for 30 min at 37°C.

First-strand cDNA synthesis

Reverse transcription was carried out by mixing the purified fragmented RNA with either 1 µl of 10 µM or 1 µl of 50 µM reverse transcription primer for random priming (*E. coli* libraries; Supplementary Table 1) or by mixing polyadenylated RNA with 4 µl 10 mM poly-dT primer (miRNA libraries; Supplementary Table 1). Reverse transcription primer and RNA mixtures were denatured by incubation at 65°C for 5 min before placing on ice. Reverse transcription was then carried out using PrimeScript Reverse Transcriptase (Takara Bio, cat. # 2680A) in 20 µl reactions of 1× reverse transcription buffer and 0.5 mM dNTP. For *E. coli* libraries, 0.5 M sorbitol and 0.1 M trehalose was added and samples were incubated at 25°C for 30 sec, 30°C for 10 min, 42°C for 30 min, 50°C for 10 min, 56°C for 10 min, 60°C for 10 min before placing on ice. For miRNA libraries, the samples were incubated 30°C for 10 min, 42°C for 60 min, 70°C for 10 min before placing on ice.

Exo I treatment

Purified reverse transcription reactions were digested with 20 U Exo I (New England Biolabs, cat. # M0293) in a total volume of 20 µl in 1× Exo I buffer or in 1× reverse transcription buffer, for 30 min (miRNA libraries) or 45 min (*E. coli* libraries) at 37°C.

E. coli libraries that did not undergo Exo I treatment were mock-treated by adding H₂O instead of enzyme and using the same conditions as Exo I-treated *E. coli* libraries.

3' cDNA adapter ligation

An adapter was ligated to the cDNA 3' end by mixing all of the purified Exo I-treated or non-treated libraries with 0.5 µl 100 µM 3' cDNA adapter (Supplementary Table 1) in 1× CircLigase buffer, 50 µM ATP, 2.5 mM MnCl₂, 10% PEG6000 and 50 U CircLigase ssDNA Ligase (Epicentre, cat. # CL4115K) and for *E. coli* libraries, 2 µl 5 M betaine was added. All reactions had a final volume of 10 µl and were incubated 2 h at 60°C, 1 h at 68°C, and 10 min at 80°C.

Library amplification by PCR

For *E. coli* libraries, PCR amplification was carried out using all of the ligated cDNA material, whereas 1/20 volume of the ligated cDNA was used for the miRNA libraries.

Reactions contained 0.6 µM forward PCR primer (Supplementary Table 1), 0.5 µM reverse PCR index primer (Supplementary Table 1), 1× Phusion HF buffer, 0.2 mM dNTPs and either 1 U (miRNA libraries) or 2 U (*E. coli* libraries) Phusion® High-Fidelity DNA Polymerase (New England Biolabs, cat. # M0530). The PCR reactions were denatured for 3 min at 98°C, followed by four cycles of 98°C for 80 s, 64°C for 15 s and 72°C for 1 min, and finally two (*E. coli* libraries) or nine (miRNA libraries) cycles of 98°C for 80 s and 72°C for 1 min, and finally 72°C for 5 min.

Sequencing

Samples were sequenced using the Illumina MiSeq (*E. coli* libraries) or NextSeq (miRNA libraries) sequencer with 2 × 75 bp paired end or 75 bp single end read lengths, respectively. The sequencing data are available at [26]. A detailed description of the data analysis is available in the Supplementary methods. ►



Blot-to-blot consistency. Time after time.

To achieve meaningful comparison of expressed proteins from blot-to-blot, trustworthy data is essential. **BlotCycler™** is designed to eliminate slight variations in fluid delivery, timing and washing that may occur when protocols are performed manually. With rigorous precision, **BlotCycler** delivers an affordable solution to busy labs demanding reproducible results.



BlotCycler™
by PRECISION BIOSYSTEMS

www.blotcycler.com

TRY OUR FREE 1-MONTH TRIAL | CALL 1-888-490-4443

► ³²P end labeling of primer & denaturing acrylamide gel electrophoresis

Primer degradation and trimming of cDNA by Exo I were assessed by gel electrophoresis of radioactively labeled cDNA and primers. For ³²P 5' end labeling of primer, 20 pmol reverse transcription poly_dT primer (Supplementary Table 1) was labeled with 100 pmol [γ -³²P] ATP (6000 Ci/mmol) by 10 U T4 Polynucleotide kinase (New England Biolabs, cat. # M0201) in 1× T4 PNK buffer for 30 min at 37°C. Labeling was terminated by adding ethylenediaminetetraacetic acid to a final concentration of 25 mM and incubating at 75°C for 10 min prior to purification by Oligo Clean & Concentrator™ (Zymo Research, cat. # D4061). Polyadenylated miRNAs were reverse transcribed with ³²P poly_dT primer and either treated with Exo I or not, as described above. Samples were then run on a 7 M urea 16% acrylamide gel.

RESULTS & DISCUSSION

Exo I treatment reduces primer–adapter contamination

Exo I selectively catalyzes the degradation of single-stranded DNA in the 3' to 5' direction, whereas double-stranded DNA is protected from degradation. In 3' cDNA ligation-based sequencing library preparation, the contamination with primer–adapter products is a recurrent issue [13]. We reasoned that treatment with Exo I could potentially degrade excess reverse transcription primers without affecting the cDNAs in complex with the template RNA (Figure 1). To test this strategy, we prepared five different sequencing libraries, each from the same amount of *E. coli* total RNA, using random primed reverse transcription and 3' cDNA ligation. The library prepared using the standard protocol shows a prominent contamination with a 147-bp PCR product, which corresponds with the expected length of the adapter being ligated directly to the reverse transcription-primer and PCR amplified in the subsequent PCR (Figure 2A). In addition, we observe a broad peak of approximately 210 bp length, which predominantly consists of longer PCR concatemers of the primer–adapter byproduct (Figure 2A). Some of these byproducts contain the Illumina adapter sequences and will therefore reduce the number of informative reads obtained by sequencing. Neither reduced reverse transcription-primer

concentration (Figure 2B) nor two rounds of size selection using solid phase reversible immobilization beads (Figure 2C) significantly reduce the primer–adapter contamination in the libraries. In contrast, Exo I treatment for 45 min or Exo I treatment in combination with reduced reverse transcription-primer concentration significantly reduce the level of primer–adaptor byproducts present in the final libraries (Figure 2D & E). Sequencing and subsequent analysis of the sequence reads obtained from the different libraries confirm the qualitative interpretation from the bioanalyzer traces. The library prepared according to the standard protocol contains 75.2% primer–adapter reads, whereas reduction of reverse transcription-primer concentration or inclusion of additional bead-based cleanup result in 78.5 and 63.5% primer–adapter contamination, respectively (Figure 2F). In contrast, the libraries prepared with Exo I treatment or Exo I treatment in combination with reduced reverse transcription-primer contain 4.7 and 1.6% primer–adapter sequence reads (Figure 2F), demonstrating that RNA–cDNA duplexes are protected from Exo I, whereas the unprotected reverse transcription-primer is degraded. The removal of surplus reverse transcription primers by Exo I treatment is easy to perform as the cDNA–RNA duplexes from reverse transcription can be purified with solid phase reversible immobilization beads or phenol–chloroform extraction followed by precipitation without disrupting the duplexes.

Application of Exo I treatment to small RNA sequencing libraries

When sequencing small RNAs, such as miRNAs, primer–adapter contamination is especially problematic, because it is difficult to separate the library products from the adapter contamination products [16] and to our knowledge, no published study has so far employed 3' cDNA ligation to prepare small RNA sequencing libraries. Instead, small RNA library preparation is typically based on ligation of adapters to the 3' and 5' ends of the small RNA. Exo I treatment has the potential to alleviate the need for size selection of cDNA products in 3' cDNA ligation-based small RNA library preparation. To test this possibility, we used a mix of five chemically synthesized miRNA

controls of 22 nt in length for library preparation with or without Exo I treatment. A polyA tail was added to the miRNA controls using Poly(A) polymerase I to facilitate reverse transcription with a poly-dT primer, ending in a degenerate VN sequence to direct priming to the miRNA–polyA junction. As expected, the small RNA-seq library prepared using the standard procedure displays abundant primer–adapter contamination (Figure 3A & Supplementary Figure 1A & D). In contrast, primer–adapter contamination is less prominent in the library having an Exo I treatment step included in the library preparation protocol after cDNA ligation and prior to library amplification (Figure 3B & Supplementary Figure 1B & E). We also tested the effect of including Exo I treatment immediately after reverse transcription without changing the buffer. Again, we observe a reduction in the amount of primer–adapter contamination in the library (Figure 3C & Supplementary Figure 1C & F), although not as large as observed with Exo I treatment in the optimal buffer. Sequencing of the libraries and analysis of the resulting sequence reads demonstrate that the rate of primer–adapter contamination is 72.2 % in the library prepared with the standard protocol as compared to 5.2 and 18.9% in the two Exo I treated libraries (Figure 3D). The relative numbers of mapped sequence reads in each of the three libraries are highly similar between each of the five different control miRNAs with an approximately fourfold increase in mapped sequence reads in the Exo I-treated libraries when compared to libraries based on the standard protocol (Supplementary Figure 2). This demonstrates that Exo I treatment can increase the proportion of informative sequencing reads in small RNA sequencing libraries based on 3' cDNA ligation by degradation of the reverse transcription primer. Therefore, small RNA sequencing protocols based on 3' cDNA ligation and including Exo I treatment are an alternative to the currently used methods.

A potential risk of treatment with Exo I is that breathing of the cDNA 3' ends in the cDNA–RNA duplexes allows exonucleolytic trimming by Exo I, which would limit the potential application of Exo I treatment. We plotted the start positions of the sequence reads mapping to the five miRNAs for each

of the three libraries. The distribution of start positions is nearly identical in the Exo I treated and untreated libraries, showing that Exo I treatment does not lead to trimming of the cDNA in the cDNA–RNA duplex (Figure 3E). We also assessed extent of cDNA trimming on a denaturing acrylamide gel and found that the length distribution of the extended cDNA is nearly identical between the Exo I-treated and the control samples, whereas non-elongated reverse transcription primer is degraded by the Exo I treatment (Figure 3F). Together, these observations demonstrate that the Exo I treatment does not trim cDNA 3' ends when present in a cDNA–RNA duplex.

FUTURE PERSPECTIVE

Sequencing-based methods are becoming increasingly important for biological research and optimization of methods for preparing RNA sequencing libraries is therefore a high priority. We show that inclusion of Exo I treatment in library preparation protocols based on 3' cDNA ligation reduces the formation of primer–adapter dimers and does not lead to trimming of the cDNA 3' end. Inclusion of Exo I treatment in sequencing library preparation protocols is simple and quick and could therefore potentially improve existing methods. In fact, Exo I treatment has been successfully included into two different sequencing library preparation protocols to facilitate cDNA 3' end ligation without primer–adapter formation [24,25], although these studies did not report the consequences of including the Exo I treatment. Moreover, Exo I treatment renders 3' cDNA ligation-based protocols relevant for RNA-seq applications, in which the formation of primer–adapter byproducts otherwise makes library preparation problematic. Here, we show that Exo I treatment allows the use of 3' cDNA ligation method for small RNA libraries and greatly increases the fraction of informative sequence reads. Finally, because no cDNA 3' trimming occurs with Exo I treatment of the RNA–cDNA duplex, Exo I treatment could also be included in sequencing methods requiring precise mapping of reverse transcription termination sites.

SUPPLEMENTARY DATA

To view the supplementary data that accompany this paper please visit the journal website at: www.future-science.com/doi/suppl/10.2144/btn-2018-0178

AUTHOR CONTRIBUTIONS

CHE, AOF, LDP and JV designed experiments. CHE and AOF performed experiments and data analysis. CHE, AOF, LDP and JV wrote the manuscript.

FINANCIAL & COMPETING INTERESTS DISCLOSURE

This work is partly funded by the Innovation Fund Denmark (IFD; IFD file 5163-00010B). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

ACKNOWLEDGMENTS

The authors wish to thank Amal Al-Chaer for experimental support.

OPEN ACCESS

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

REFERENCES

- Shendure J, Balasubramanian S, Church GM *et al*. DNA sequencing at 40: past, present and future. *Nature* 550(7676), 345–353 (2017).
- Shendure J, Lieberman Aiden E. The expanding scope of DNA sequencing. *Nat. Biotechnol.* 30(11), 1084–1094 (2012).
- Head SR, Komori HK, Lamere SA *et al*. Library construction for next-generation sequencing: overviews and challenges. *BioTechniques* 56(2), 61–64, 66, 68, passim (2014).
- Bentley DR, Balasubramanian S, Swerdlow HP *et al*. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456(7218), 53–59 (2008).
- Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* 109 21 29 21–29 (2015).
- Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 30(4), 892–897 (2001).
- Routh A, Head SR, Ordoekhanian P, Johnson JE. Click-Seq: fragmentation-free next-generation sequencing via click ligation of adaptors to stochastically terminated 3'-azido cDNAs. *J. Mol. Biol.* 427(16), 2610–2616 (2015).
- Hou Z, Jiang P, Swanson SA *et al*. A cost-effective RNA sequencing protocol for large-scale gene expression studies. *Sci. Rep.* 5, 9570 (2015).
- Poulsen LD, Vinther J. RNA-Seq for bacterial gene expression. *Curr. Protoc. Nucleic Acid Chem.* 73(1), e55 (2018).
- Lucks JB, Mortimer SA, Trapnell C *et al*. Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc. Natl Acad. Sci. USA* 108(27), 11063–11068 (2011).
- Kielipinski LJ, Vinther J. Massive parallel-sequencing-based hydroxyl radical probing of RNA accessibility. *Nucleic Acids Res.* 42(8), e70 (2014).

- Ding Y, Tang Y, Kwok CK, Zhang Y, Bevilacqua PC, Assmann SM. *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* 505(7485), 696–700 (2014).
- Ritchey LE, Su Z, Tang Y, Tack DC, Assmann SM, Bevilacqua PC. Structure-seq2: sensitive and accurate genome-wide profiling of RNA structure *in vivo*. *Nucleic Acids Res.* 45(14), e135 (2017).
- Raabe CA, Tang TH, Brosius J, Rozhdetsvensky TS. Biases in small RNA deep sequencing data. *Nucleic Acids Res.* 42(3), 1414–1426 (2014).
- Watters KE, Abbott TR, Lucks JB. Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq. *Nucleic Acids Res.* 44(2), e12 (2016).
- Shore S, Henderson JM, Lebedev A *et al*. Small RNA library preparation method for next-generation sequencing using chemical modifications to prevent adapter dimer formation. *PLoS ONE* 11(11), e0167009 (2016).
- Spitale RC, Crisalli P, Flynn RA, Torre EA, Kool ET, Chang HY. RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* 9(1), 18–20 (2013).
- Kodzius R, Kojima M, Nishiyori H *et al*. CAGE: cap analysis of gene expression. *Nat. Methods* 3(3), 211–222 (2006).
- Poulsen LD, Kielipinski LJ, Salama SR, Krogh A, Vinther J. SHAPE Selection (SHAPEs) enrich for RNA structure signal in SHAPE sequencing-based probing data. *RNA* 21(5), 1042–1052 (2015).
- Lehman IR, Nussbaum AL. The deoxyribonucleases of *Escherichia coli*. V. On the specificity of exonuclease I (phosphodiesterase). *J. Biol. Chem.* 239, 2628–2636 (1964).
- Werle E, Schneider C, Renner M, Volker M, Fiehn W. Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res.* 22(20), 4354–4355 (1994).
- Tang F, Barbacioru C, Wang Y *et al*. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6(5), 377–382 (2009).
- Sasagawa Y, Nikaido I, Hayashi T *et al*. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol.* 14(4), R31 (2013).
- Van Nostrand EL, Pratt GA, Shishkin AA *et al*. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 13(6), 508–514 (2016).
- Ordoekhanian P, Nichols J, Head SR. Primer extension, capture, and on-bead cDNA ligation: an efficient maseq library prep method for determining reverse transcription termination sites. *Methods Mol. Biol.* 1712, 253–261 (2018).
- Index of /~vinther/data/Exol-seq-data/data. <http://people.binf.ku.dk/~vinther/data/Exol-seq-data/data/>